

## Reply

## 5. Subject Matter of Reply

(1) In order to distinguish the present invention from the cited documents cited in the opinion of the International Searching Authority, the claims are amended in the written amendment submitted together with this reply.

Concretely, Claim 1 is amended so as to include the restricting matters of Claim 8 before amendment. In addition, the grounds for this amendment are described in paragraphs [0042] and [0043] in the specification. Furthermore, the claim is amended so as not to use expressions "first" and "second" as in "first voice quality." Accordingly, amendment of Claim 1 does not constitute addition of new matter.

In addition, Claims 2 to 7 are amended so as to match with Claim 1 after amendment. Accordingly, amendment of these claims does not constitute addition of new matter.

In addition, Claims 11 and 18 are amended in the same manner as Claim 1. Accordingly, amendment of these claims does not constitute addition of new matter.

Furthermore, Claims 8 to 10 and Claims 12 to 17 are deleted.

(2) The invention of Claim 1 after amendment is not described in any of the cited documents cited in the opinion of the International Searching Authority (Document 1: JP9-244693A; Document 2: JP2002-351489A; Document 3: JP4-158397A).

In the invention according to Claim 1 after amendment, the designating unit (including a display) places fixed points which indicate the voice quality of a plurality of pieces of voice element information which are stored in the memory unit at, for example, two-dimensional coordinates, and places a plurality of icons (set points) on the basis of operation by the user at the above described two-dimensional coordinates. In addition, the designating unit derives the ratio of contribution of each of the above described plurality of pieces of synthetic voice information to morphing on the basis of the placement of moving points which continuously shift along the time sequence between a plurality of set points and fixed points, and designates this rate for the morphing unit. As a result, the morphing unit changes the rate of contribution of each of the plurality of pieces of synthetic voice information to intermediate synthetic voice information along the time sequence.

As described above, in the invention according to Claim 1 after amendment, the designating unit changes the ratio of contribution of voice qualities that are different from each other to morphing along the time sequence in accordance with operation by the user, and therefore, a voice synthesis device having high expressiveness; for

example "cool at the beginning of speech and gradually getting angry while speaking," which was conventionally impossible, can be implemented (paragraphs [0059] and [0060]).

That is to say, in the invention according to Claim 1 after amendment, voice quality can be gradually and chronologically changed between two voice qualities, and in addition, many different voice qualities can be expressed on a display, and arbitrary set points of a plurality of voice qualities on a display can be designated by the user, and therefore, chronological voice morphing with great variety becomes possible.

FIG. 4 of the present application shows the effects of the invention according to Claim 1 after amendment. In FIG. 4, the voice quality of the synthetic voice that is outputted for the first time is close to voice quality Z because designating icon (moving point) 104i is at icon (moving point) 21 that is close to voice quality Z. However, designating icon 104i gradually moves to icon (set point) 22 that is close to voice quality A, and after that, moves away from voice qualities A and Z to icon (set point) 23 that is close to voice quality B. Accordingly, the voice quality of the synthetic voice gradually changes in such a manner as to become close to voice quality A from the voice quality which is close to voice quality Z, and after that, close to voice quality B.

As can be seen from this figure, in the invention according to Claim 1 after amendment in the present application, it is possible to designate voice quality with extremely great variety in such a manner that the voice quality moves away from a plurality of qualities (voice qualities Z and A) and closer to another voice quality (voice quality B), in addition to designating the voice quality between two voice qualities.

Accordingly, the user can confirm the voice quality of a plurality of pieces of voice element information stored in a memory unit as fixed points on a display and designate set points for chronologically changing the voice quality on the display, and therefore, speech morphing between a great plurality of different voice qualities can be intuitively operated in a variety of manners.

Therefore, we firmly believe that the present invention should not be denied patentability on the basis of the cited documents.

## 答 弁 書

特許庁審査官 殿

1. 国際出願の表示 PCT/J P 2005/000505

### 2. 出願人

名称 松下電器産業株式会社  
MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.  
あて名 〒571-8501 日本国大阪府門真市大字門真 1 0 0 6 番地  
1006, Oaza Kadoma, Kadoma-shi, Osaka 571-8501 Japan  
国籍 日本国 Japan  
住所 日本国 Japan

### 3. 代理人

氏名 (10921) 弁理士 新居広守  
NII Hiromori  
あて名 〒532-0011 日本国大阪府大阪市淀川区西中島 3 丁目 1 1 番 2 6 号  
新大阪末広センタービル 3 F 新居国際特許事務所内  
c/o NII Patent Firm, 3<sup>rd</sup> Floor, Shin-Osaka Suehiro Center Bldg.,  
11-26, Nishinakajima 3-chome, Yodogawa-ku, Osaka-shi, Osaka  
532-0011 JAPAN



4. 通知の日付 10.05.2005

### 5. 答弁の内容

(1) 国際調査機関の見解書で挙げられた引用文献と本願発明を区別するために、この答弁書と同時に提出する手続補正書にて、請求の範囲の補正を行いました。

具体的に、請求項 1 については、補正前の請求項 8 の限定事項等を含めるように補正しました。また、このような補正の根拠は、明細書の [0042] 段落および [0043] 段落に記載されております。さらに、「第 1 の声質」のような「第 1」および「第 2」という表現を使用しないように改めました。したがって、請求項 1 の補正は新規事項を追加するものではありません。

また、請求項 2～7 については、補正後の請求項 1 に整合するように補正しました。した

がって、これらの請求項の補正は新規事項を追加するものではありません。

また、請求項 1 1 および請求項 1 8 については、請求項 1 と同様の補正を行いました。したがって、これらの請求項の補正は新規事項を追加するものではありません。

さらに、請求項 8 ～ 1 0 と請求項 1 2 ～ 1 7 を削除しました。

(2) 補正後の請求項 1 の発明は、国際調査機関の見解書で挙げられた引用文献（文献 1 : JP9-244693A、文献 2 : JP2002-351489A、文献 3 : JP4-158397A）の何れにも記載されておられません。

補正後の請求項 1 に記載の発明では、指定手段（ディスプレイを含む）は、記憶手段に記憶されている複数の音声素片情報の声質を示す固定点を、例えば二次元座標上に配置するとともに、ユーザの操作に基づいて複数のアイコン（設定点）を上記二次元座標上に配置します。そして、指定手段は、複数の設定点間を時系列に沿って連続的に移動する移動点および固定点の配置に基づいて、モーフィングに寄与する前記複数の合成音声情報のそれぞれ割合を導出し、その割合をモーフィング手段に指定します。その結果、モーフィング手段は、複数の合成音声情報の中間合成音声情報に寄与するそれぞれの割合を時系列に沿って変化させます。

このように、補正後の請求項 1 に記載の発明では、ユーザによる操作に応じて、互いに異なる声質のモーフィングに寄与する割合を指定手段が時系列に沿って変化させるため、「喋り始めは冷静だが、喋りながら段々と怒っていく」というような、従来は不可能だった表現力の高い音声合成装置を実現することができます（〔0059〕段落及び〔0060〕段落）。

即ち、補正後の請求項 1 に記載の発明では、2つの声質間を徐々に時間的に変化させるだけでなく、ディスプレイ上に多くの異なる声質を表現でき、且つ、ディスプレイ上で複数の声質の設定点をユーザが任意に指定できるので、バラエティに富んだ時系列的な音声モーフィングが可能となります。

本願の図 4 では、補正後の請求項 1 に記載の発明の効果を示すものです。図 4 では、最初に出力される合成音声の声質は、指定アイコン（移動点）104i が声質 Z に近いアイコン（設定点）21にあるので、声質 Z に近いです。しかし、指定アイコン 104i は、徐々に声質 A に近いアイコン（設定点）22に移動し、その後、声質 A および声質 Z から離れて、声質 B に近いアイコン（設定点）23に移動します。したがって、合成音声の声質は、声質 Z に近い声質から声質 A に近づいて、その後、声質 B に近づくように次第に変化します。

この図からもわかるように、本願の補正後の請求項 1 の発明では、2つの声質間の指定だけでなく、複数の声質（声質 Z および声質 A）から遠ざけて他の声質（声質 B）に近づけるといった非常にバラエティに富んだ声質の指定も可能となります。

したがって、ユーザは、記憶手段で記憶されている複数の音声素片情報の声質をディスプレイ上で固定点として確認でき、かつ、ディスプレイ上で時間的に声質変化させるための設

定点を指定できるので、異なる多数の声質間における音声モーフィングを、直感的に、かつ、多様に操作することができます。

よって、本願発明の特許性は、引用文献に基づいて否定されないものと確信致します。

以上